PSYCHOLOGICAL SCIENCE

Research Report

WHERE WE GO WITH A LITTLE GOOD INFORMATION

Ranxiao Frances Wang and James E. Cutting
Cornell University

Abstract—*When observers move through an environment, they are immersed in a sea of motions that guide their further movements. The horizontal relative motions of all possible pairs of stationary objects fall into three classes: They converge, diverge and slow down, or diverge with increasing velocity. Conjoined with ordinal depth information, the first two motions reveal nominal invariants, constraining heading to one side of the visual field. When two object pairs yield invariants on opposing sides of the heading, they can constrain judgments to a narrow region. Distributional analyses of responses in an experiment involving simulated observer movement suggest that observers follow these constraints.*

When people walk, run, or drive through the world, how do they know where they are going? How do they know their heading so they can safely avoid obstacles in their path? Over the past 50 years, particularly because of the work of Gibson (1950, 1979), these questions have received sustained attention. Over the past 25 years, advances in neurophysiology have revealed cells in the visual systems of monkeys, pigeons, and cats that respond to relative motion pooled over wide regions of the visual field (Allman, Miezin, & McGuinness, 1985; Bradley, Maxwell, Andersen, Banks, & Shenoy, 1996; Bridgeman, 1972; Frost & Nakayama, 1983; Pasternak, Albano, & Harvitt, 1990). Thus, many researchers have sought computational and psychophysical evidence that human beings might negotiate environments on the basis of such cells (Hildreth, 1992; Nakayama & Loomis, 1974; Rieger & Lawton, 1985; Warren & Saunders, 1995). The supporting evidence is substantial, provided that test environments contain many elements. Motion pooling fails, however, when simulated navigation occurs through relatively sparse environments (Cutting, 1996). In such environments, however, considerable accuracy is achieved by human observers from the relative motions of a few pairs of stationary objects. Our data suggest that this accuracy is based on multiple constraints derived from relative motions of these object pairs. Accurate heading judgments in natural environments can also be accomplished in this way.

Mathematically, the motions of stationary objects around a moving observer can be parsed in several ways (see Cutting, 1986; Cutting, Springer, Braren, & Johnson, 1992; Koenderink & van Doorn, 1975; Longuet-Higgins & Prazdny, 1980; Regan & Beverley, 1982). Here we develop a new approach, considering the relative motions of pairs of stationary objects with respect to the eye of the moving observer. The horizontal motion of all possible pairs falls into three classes: Object pairs can converge, they can diverge and slow down, or they can diverge with increasing velocity (Cutting, 1996). In the forward visual field, convergence is always acceleratory, except in certain cases when the observer is moving along a curved path. Figure 1 shows these three classes of motion with respect to an object to the right of one's path. Two important rules about heading, also shown in Figure 1, then follow. First, when objects converge, one's instantaneous heading is always outside of the nearest member of the pair, in this case to the left. This rule has no exceptions. Second, when objects decelerate apart, the same is true. This rule is qualified in that both objects must be within 45° of one's heading, but this condition is not overly restrictive because about 90% of pedestrian gazes fall within this bound (Wagner, Baird, & Barbaresi, 1981). These two rules are optical invariants, or statistically certain sources of information (Gibson, 1979).

There is also a third relative-motion class. When a pair of objects accelerate apart, one's heading is unsure. To compute the efficacy of accelerating divergence as it predicts one's heading, we assumed that a pedestrian's gaze is within ±90° of the heading on a reference object at 30 m, near the median distance for pedestrian fixation (Wagner et al., 1981), and that the second object under consideration is between 1 and 100 m and within ±20° to either side of the first. We then sampled gaze-heading angles at 1° intervals between ±90°, computed the relative areas of the three region types shown in Figure 1, and then weighted angular gaze-heading calculations by their naturally occurring frequency (Wagner et al., 1981). The resulting values are shown at the bottom of Figure 1. (If the ratio of depths is known, further refinements can be made.) Thus, this motion class yields an optical heuristic, a probabilistic information source (Gilden & Proffitt, 1989), suggesting that heading is most often to the outside of the farther object in the pair.

Notice that, taken singly, none of these potential sources of information predicts one's absolute heading, or the exact direction in which one is moving. In particular, the two invariants nominally constrain one's heading direction; they specify that it is left or right of a particular object, but not by how much. Yet there is ample evidence that observers can report their absolute headings with reasonable accuracy (Royden, Banks, & Crowell, 1992; van den Berg & Brenner, 1994; Warren & Hannon, 1988). How might nominal invariants yield accurate, near-metric heading judgments? Our answer is that invariants on both sides of one's heading constrain judgments to a narrow region.

METHOD

We simulated observer movement through minimal, but moderately naturalistic, environments with only two, three, four, and seven schematic trees. Sample final frames are shown in Figure 2, column a. With two trees, heading may fall into one of three categories—left, between, and right—labeled 1, 2, and 3 in Figure 2, column b. We can generalize that the number of trees (N) creates $N + 1$ heading categories. These columns in the figure also show the mean separations between adjacent pairs of trees in the experiment. We generated three types of trials—those with heading *unspecified* and diverging acceleration in all pairs of trees, those with heading *nominally specified* by

---

Ranxiao Frances Wang is now at the Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology.

Address correspondence to James Cutting, Department of Psychology, Uris Hall, Cornell University, Ithaca, NY 14853-7601; e-mail: jec7@cornell.edu.

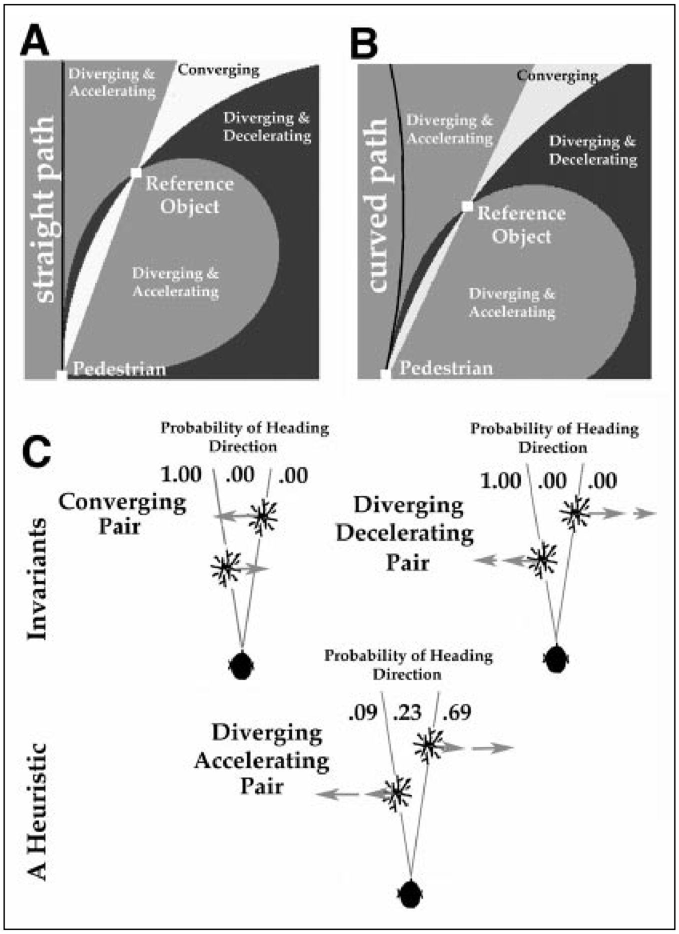

Fig. 1. Plan views of a pedestrian on a linear (a) and a circular (b) path and the nominal rules for heading judgments (c). In (a) and (b), a pedestrian is looking off his or her path at a particular reference object, one potential member of an object pair. Any second object in the region marked "converging" will move toward that reference object; any second object in the region marked "diverging and decelerating" will move away from the reference object at a decreasing velocity; and any second object in the region marked "diverging and accelerating" will move away from the reference object with increasing velocity. Different placements of the reference object will alter the layout of these regions, but the patterns will generally remain. Of the three rules (c) that emerge from the consideration of such object pairs, two are invariant rules that always specify that heading is to the outside of the near member of the object pair, and one is a heuristic rule that suggests that heading is most often to the outside of the far member of the pair.

convergence or diverging deceleration in at least one pair of trees, and those with heading *categorically specified* by pairs of invariants to either side of the heading. Our analysis depends on observers' ability to discern the ordinal depth of trees in the display. The simulated distance of each tree from the observer was given by relative size and height in the visual field, both powerful sources of ordinal depth information (Cutting & Vishton, 1995).

Motion sequences were generated at 17 frames/s on a Silicon Graphics Iris Workstation (Model 4D/35GT), but frame rates are known to have little effect on heading judgments in these situations (Vishton & Cutting, 1995). Viewers sat 0.5 m from the monitor, yielding 40°-wide displays at a resolution of about 30 pixels/°. Sequences simulated 4 s of observer movement at 1.25 eye heights/s, while fixating on a tree at midscreen whose initial distance was 14.7 eye heights. Simulated pursuit-fixation sequences were used to test the adequacy of visual information to the heading-judgment task without feedback from eye movements. Sequences with two, three, and four trees were generated along a linear path; those with seven trees were generated along a circular path (radius = 150 eye heights). Final gaze-heading angles were 3.1° and 6.2° for linear paths and 2.3° and 5.6° for circular paths.

Ten naive observers viewed 6 to 12 practice trials with nominal feedback, then 192 linear-path sequences without feedback: 8 different motions patterns (generally 6 with convergence, decelerating divergence, or both and 2 with accelerating divergence in all pairs) × 3 different numbers of trees (two, three, and four) × 2 gaze-heading angles × 2 heading directions (left and right) × 2 replications. Among the two-tree sequences, 24 were unconstrained, and 40 constrained responses to Heading Category 1 (see Fig. 2, column b). Among the three-tree sequences, 16 were unconstrained, 16 constrained responses to Heading Category 1, and 32 constrained responses to Heading Categories 1 and 2. Among the four-tree sequences, 16 were unconstrained, 8 constrained responses to Heading Category 1, 24 constrained them to Heading Categories 1 and 2, 12 constrained them to Heading Categories 1 through 3, and 4 constrained them to Heading Category 3 only. The same observers then viewed 56 circular-path sequences with seven trees: 7 motion configurations × 2 heading directions × 2 paths (curving left or right) × 2 replications. None of the sequences were unconstrained; 8 and 32 constrained responses to Heading Categories 1 through 4 and 1 through 3, respectively; and 16 constrained them to Heading Category 3 only.

At the end of each trial, motion ceased and the final frame remained on the screen. Observers moved the mouse-controlled screen cursor to their perceived heading at the horizon and clicked a mouse button, which started the next trial.

RESULTS

At simulated velocities used here, heading judgments should be within ±3.7° of the instantaneous heading (Cutting et al., 1992; Vishton & Cutting, 1995). Overall, mean error was 2.4°. Absolute responses were then divided into heading categories. For illustrative purposes, half of the stimuli and their responses were flipped around array midlines to show true heading always to the left. When heading was unspecified, responses were generally distributed around the middle categories, as shown in Figure 2, column c. Responses were never greater than 50% in any category, and mean error was 4.1°. However, when heading was nominally specified, responses were almost always placed in the category next to the exclusion boundary (82%), as shown in Figure 2, column d; mean error was 2.2°. Most important, some four- and seven-tree sequences excluded all but one heading category, and judgments were extremely accurate (81%), as shown in Figure 2, column e; mean error was 0.8°. Notice the striking increase in absolute accuracy with increasing constraints, $F(2, 18) = 41.1$, $p < .0001$. Dark regions in the graphs indicate those categories for which heading is excluded by the nominal invariants or by the heuristic given known distances. Across all trials with heading nominally or categorically specified, 65% of all possible response categories were excluded by the nominal invariants or by the heuristic, yet only 16% of all responses occurred within these categories.

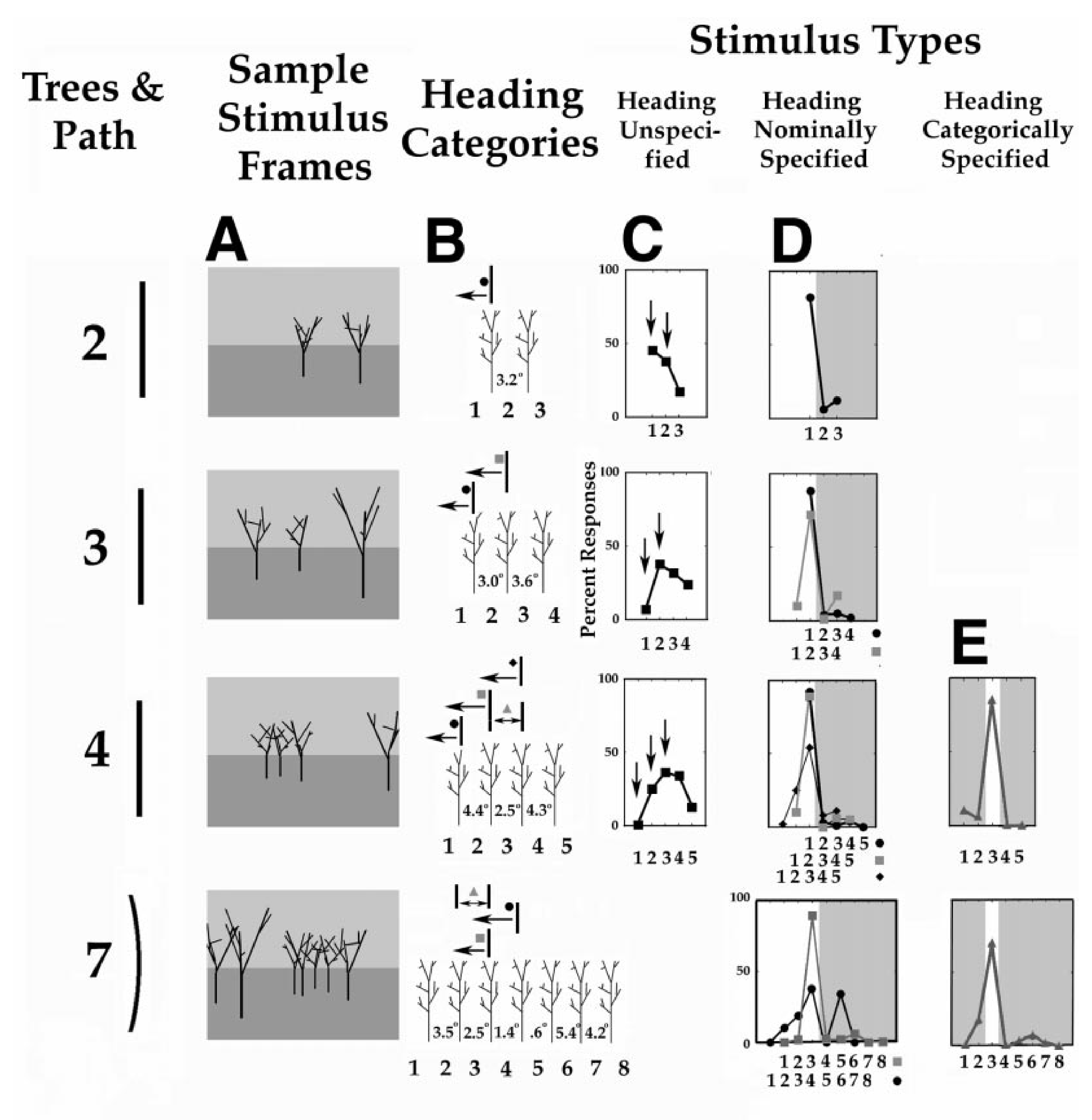

Fig. 2. Stimuli and results of the experiment. For stimuli with two through four trees, paths were always linear; for those with seven trees, the path was always circular. Column a shows sample final frames from the various stimuli. Column b shows the mean separations between the trees in the various stimuli, the numbering of the heading categories from left to right, and the nominal and categorical constraints on the various types of trials. Column c shows the results for those trials with heading unspecified (i.e., with all tree pairs diverging and accelerating). Arrows indicate categories with true headings within them. Column d shows the results for those trials with heading nominally specified; categorical results are aligned to the exclusion boundary, the near tree of the invariant pair in each stimulus. Column e shows the results for trials with two invariants of opposite sign, and thus having heading categorically specified. Dark regions indicate those categories where responses should not occur according to the invariant rules. Results and constraints are shown as if the correct heading were always to the left of midscreen.

What is the evidence that observers' responses reflect categorical rather than absolute heading information? Computational measures that pool motions to estimate absolute heading do not fare well with these stimuli. One adds each motion vector in the visual field (Warren & Saunders, 1995), and another squares vectors before summation (Rieger & Lawton, 1985). When either is applied to motions in a more cluttered environment, a difference vector from these sums generally points to the heading, and its magnitude generally indicates how far the observer is from the moving objects (Cutting, 1996). In the relatively uncluttered environments tested in the present study, however, neither of these last two measures correlated with observers' absolute responses (*r*s < .14, *p*s > .17), and neither predicted heading direction from the center of the display (*r*s < .06, *p*s > .38).

If observers' heading judgments have only categorical guidance, certain distributional characteristics should follow. First, consider selected data from the two- through four-tree sequences, shown in columns c and d of Figure 2. One would expect each viewer's absolute responses to trials with unspecified headings to have more variance and less skew than those to trials with nominally specified headings. This expectation was confirmed: Mean standard deviations were 3.2° and 2.6°, respectively, $F(1, 9) = 8.3$, $p < .02$, and mean skews were 0.42° and 0.90°, respectively, $F(1, 9) = 6.0$, $p < .04$. Second, consider the data from selected four- and seven-tree sequences, shown in columns d and e of Figure 2. One would expect response variance and skews for trials with nominally specified headings to be greater than those for trials with categorically specified headings. This pattern, too, occurred: Mean standard deviations were 2.5° and 1.4°, $F(1, 9) = 5.5$, $p < .04$, and mean skews were 1.22° and 0.15°, $F(1, 9) = 5.37$, $p < .05$. In addition, consider all trials with heading nominally specified. Although judgments clustered near the exclusion boundary, that category did not always contain the true heading. When this dissociation occurred, responses were more frequent in the boundary category (73%) than in its neighbor with the true heading (10%), $t(10) = 10.3$, $p < .0001$. Thus, heading judgments appear to be distributed across categories according to the bounds given by the nominal invariants. No pooling scheme seems likely to be able to predict such patterns.

DISCUSSION

Three questions may arise about our analysis and how it relates to other approaches to perception. First, given that accelerations and decelerations are not particularly easy to perceive, can they serve as a basis for heading judgments? Our answer is twofold. On the one hand, convergence was the more potent invariant in these data. Mean error for stimuli with only convergence was 2.0°, whereas that for stimuli with only diverging deceleration was 2.7°, $F(1, 9) = 6.7$, $p < .03$. On the other hand, psychophysical evidence shows that decelerations are easier to detect than accelerations, and over a 1-s period, it is relatively easy to perceive a drop in velocity to 42% of the initial value (Schmerler, 1976). On those trials in which deceleration was the only information available to constrain heading, the median velocity drop was to 39% over the last second of the trial. Thus, the decelerations generated in the present study were generally perceptible.

Second, in light of our results, what should one conclude from previous neurophysiological findings on pooled motions? In all such studies thus far, fields of moving dots have been used. These are known to yield occasionally different results than stimuli with more naturalistic layout (Cutting, Vishton, Flückiger, Baumberger, & Gerndt, 1997; Vishton & Cutting, 1995). The reason may be that it is difficult to pay attention to single elements in dot-field displays. In our displays, however, trees can easily serve as foci of attention. Because it is known that attention dramatically modulates the responses of cortical cells (e.g., Motter, 1993), we suggest that attention, when it can, would also modulate the responses of cells involved in heading judgments. The cells identified in previous research are still possible candidates for the registration of the information identified here.

Third, how do these results fit into the fabric of perception in general, and ecological research in particular? On the one hand, they do not support the idea that locomotion is guided by a focus of expansion (or of radial outflow), Gibson's (1979) proposed invariant, nor do they support an analysis in terms of a "melon-shaped family of curves" (Gibson, 1979, p. 227). Moreover, they do not fit within any scheme that has generally followed Gibson (e.g., Kim, Growney, & Turvey, 1996; Warren & Hannon, 1988). On the other hand, the nominal invariants discussed here are no less a species of invariant than any other; they simply measure the world in a different way. Gibson proclaimed invariants to be the basis of all perception (but see Cutting, 1993, for doubts about their ubiquity). Insofar as this research has elucidated two new invariants, it is consistent with Gibson's ecological approach to visual perception. Insofar as it has elucidated a heuristic that may also prove useful, it is also allied to other approaches, such as that of Brunswik (1956).

CONCLUSION

Human observers can make highly accurate heading judgments on the basis of a little good information. This information arises from the nominal invariants of convergence and decelerating divergence in pairs of stationary objects. When such sources are present on both sides of one's heading, accuracy of judgments can greatly exceed situational demands. Models of heading judgments based on the pooling of motions over relatively large regions of the visual field do not account for these results.

Acknowledgments—This work was supported by Grant ASC-9523483 from the National Science Foundation. A preliminary version of this report was presented at the May 1997 conference of the Association for Research in Vision and Ophthalmology, in Ft. Lauderdale, Florida. We thank M. Flückiger, S.H. Johnson, N.E. Karwan, and P.M. Vishton for comments, and J.T. Todd for the suggestion about measuring accelerations.

REFERENCES

Allman, J., Miezin, F., & McGuinness, E. (1985). Direction- and velocity-specific response from beyond the classical receptive field in the middle temporal visual area (MT). *Perception*, *14*, 105–126.

Bradley, D.C., Maxwell, M., Andersen, R.A., Banks, M.S., & Shenoy, K.V. (1996). Mechanisms of heading perception in primate visual cortex. *Science*, *273*, 1544–1547.

Bridgeman, B. (1972). Visual receptive fields sensitive to absolute and relation motion during tracking. *Science*, *178*, 1106–1108.

Brunswik, E. (1956). *Perception and the representative design of psychological experiments*. Berkeley: University of California Press.

Cutting, J.E. (1986). *Perception with an eye for motion*. Cambridge, MA: MIT Press.

Cutting, J.E. (1993). Perceptual artifacts and phenomena: Gibson's role in the 20th century. In S. Masin (Ed.), *Foundations of perceptual theory* (pp. 231–260). Amsterdam: Elsevier Scientific.

Cutting, J.E. (1996). Wayfinding from multiple sources of local information in retinal flow. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 1299–1313.

Cutting, J.E., Springer, K., Braren, P.A., & Johnson, S.H. (1992). Wayfinding on foot from information in retinal, not optical, flow. *Journal of Experimental Psychology: General*, *121*, 41–72, 129.

Cutting, J.E., & Vishton, P.M. (1995). Perceiving layout and knowing distances. In W. Epstein & S. Rogers (Eds.), *Perception of space and motion* (pp. 69–117). San Diego: Academic Press.

Cutting, J.E., Vishton, P.M., Flückiger, M., Baumberger, B., & Gerndt, J.D. (1997). Heading and path information from retinal flow in naturalistic environments. *Perception & Psychophysics*, *59*, 426–441.

Frost, B.J., & Nakayama, K. (1983). Single visual neurons code opposing motion independent of direction. *Science*, *220*, 774–775.

Gibson, J.J. (1950). *Perception of the visual world*. Boston: Houghton Mifflin.

Gibson, J.J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.

Gilden, D.L., & Proffitt, D.R. (1989). Understanding collision dynamics. *Journal of Experimental Psychology: Human Perception and Performance*, *15*, 372–383.

Hildreth, E.C. (1992). Recovering heading for visually-guided navigation. *Vision Research*, *32*, 1177–1192.

Kim, N.-G., Growney, R., & Turvey, M.T. (1996). Optical flow not retinal flow is the basis of wayfinding on foot. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 1279–1288.

Koenderink, J.J., & van Doorn, A.J. (1975). Invariant properties of the motion parallax field due to the movement of rigid bodies relative to an observer. *Optica Acta*, *22*, 773–791.

Longuet-Higgins, H.C., & Prazdny, K. (1980). The interpretation of a moving retinal image. *Proceedings of the Royal Society of London, B*, *208*, 385–397.

Motter, B.C. (1993). Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli. *Journal of Neurophysiology*, *70*, 909–919.

Nakayama, K., & Loomis, J.M. (1974). Optical velocity patterns, velocity sensitive neurons, and space perception: A hypothesis. *Perception*, *3*, 63–80.

Pasternak, T., Albano, J.E., & Harvitt, D.M. (1990). The role of directionally selective neurons in the perception of global motion. *Journal of Neuroscience*, *10*, 3079–3086.

Regan, D.M., & Beverley, K.I. (1982). How do we avoid confounding the direction we are looking with the direction we are going? *Science*, *215*, 194–196.

Rieger, J.H., & Lawton, D.T. (1985). Processing differential image motion. *Journal of the Optical Society of America, A*, *2*, 354–360.

Royden, C.S., Banks, M.S., & Crowell, J.A. (1992). The perception of heading during eye movements. *Nature*, *360*, 583–585.

Schmerler, J. (1976). The visual perception of accelerated motion. *Perception*, *5*, 167–185.

van den Berg, A.V., & Brenner, E. (1994). Why two eyes are better than one for judgments of heading. *Nature*, *371*, 700–702.

Vishton, P.M., & Cutting, J.E. (1995). Wayfinding, displacements, and mental maps: Velocity fields are not typically used to determine one's aimpoint. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 978–995.

Wagner, M., Baird, J.C., & Barbaresi, W. (1981). The locus of environmental attention. *Journal of Environmental Psychology*, *1*, 195–201.

Warren, W.H., & Hannon, D.J. (1988). Direction of self-motion is perceived from optical flow. *Nature*, *336*, 162–163.

Warren, W.H., & Saunders, J.A. (1995). Perceiving heading in the presence of moving objects. *Perception*, *24*, 315–331.

(RECEIVED 11/3/97; ACCEPTED 1/20/98)